CompCytogen 11(4): 821–832 (2017)
doi: 10.3897/CompCytogen.v11i4.19953
http://compcytogen.pensoft.net

RESEARCH ARTICLE

COMPARATIVE Cytogenetics
A peer-reviewed open-access journal
International Journal of Plant & Animal Cytogenetics, Karyosystematics, and Molecular Systematics

Chromosomal assignment of centromere-specific histone CENH3 genes in rye (*Secale cereale* L.) and their phylogeny

Yulia A. Lipikhina[1,*], Elena V. Evtushenko[1,*], Evgeny A. Elisafenko[2], Alexander V. Vershinin[1,*]

1 *Institute of Molecular and Cellular Biology SB RAS, Novosibirsk, 630090, Russia* **2** *Institute of Cytology and Genetics SB RAS, Novosibirsk, 630090, Russia*

Corresponding author: *Alexander V. Vershinin* (avershin@mcb.nsc.ru)

Academic editor: *M. Guerra* | Received 1 August 2017 | Accepted 6 November 2017 | Published 14 December 2017

http://zoobank.org/B131CC9E-3F0F-4E1C-BE89-1FA5D1CCC6F5

Citation: Lipikhina YA, Evtushenko EV, Elisafenko EA, Vershinin AV (2017) Chromosomal assignment of centromere-specific histone CENH3 genes in rye (*Secale cereale* L.) and their phylogeny. Comparative Cytogenetics 11(4): 821–832. https://doi.org/10.3897/CompCytogen.v11i4.19953

Abstract

Centromeres are essential for correct chromosome segregation during cell division and are determined by the presence of centromere-specific histone 3 (CENH3). Most of the diploid plant species, in which the structure and copy number of *CENH3* genes have been determined, have this gene as a singleton; however, some cereal species in the tribe Triticeae have been found to have CENH3 in two variants. In this work, using the set of the wheat-rye addition lines we wanted to establish the chromosomal assignment of the *CENH3* genes in the cultivated rye, *Secale cereale* (Linnaeus, 1753), in order to expand our knowledge about synteny conservation in the most important cereal species and about their chromosome evolution. To this end, we have also analyzed data in available genome sequencing databases. As a result, the αCENH3 and βCENH3 forms have been assigned to rye chromosomes 1R and 6R: specifically, the commonest variants αCENH3-v1 and βCENH3-v1 to chromosome 1R, and the rare variants, αCENH3-v2 and probably βCENH3-v2, to chromosome 6R. No other CENH3 variants have been found by analysis of the rye genome sequencing databases. Our chromosomal assignment of CENH3 in rye has been found to be the same as that in barley, suggesting that both main forms of CENH3 appeared in a Triticeae species before the barley and wheat-rye lineages split.

Keywords

Centromeric histone CENH3, rye, wheat-rye addition lines, barley, Triticeae

* these authors contributed equally to the paper

Copyright *Yulia A. Lipikhina et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

In centromeric nucleosomes, canonical histone H3 appears in the form of its centromere-specific modification denoted in plants as CENH3. The presence of this protein is by far the most distinct molecular feature of this chromosomal region. Unlike canonical histone H3, which has a conserved structure, CENH3 normally shows considerable variability across species (Sanei et al. 2011, Maheshwari et al. 2015, Neumann et al. 2015). Different domains of this molecule are diverging differently. An extended N-terminal tail (NTT) domain and loop 1 of the histone fold domain (HFD) putatively interact with centromeric DNA (Vermaak et al. 2002) and show signatures of positive selection in some animal and plant species (Malik and Henikoff 2001, Talbert et al. 2002), while the part of the HFD that lies outside loop 1 is generally conserved (Talbert et al. 2002, Cooper and Henikoff 2004, Hirsch et al. 2009, Finseth et al. 2015).

Most of the diploid plant species, in which the structure and copy number of *CENH3* have been determined, have this gene as a singleton. Cereal species as these are, for example, maize and rice (Zhong et al. 2002, Nagaki et al. 2004), which are phylogenetically distant from the tribe Triticeae (the subfamily Pooideae, Dumort, 1824), which includes the closely related genera *Triticum* Linnaeus, 1753, *Secale* Linnaeus, 1753 and *Hordeum* Linnaeus, 1753. The closest relatives to rye in Triticeae, namely, tetraploid wheat species (Yuan et al. 2015), diploid barley, wheat and *Aegilops* Linnaeus, 1753 species (Sanei et al. 2011, Yuan et al. 2015), have been found to have *CENH3* in two variants. Thus, the presence of at least two copies of the *CENH3* gene is a shared feature of the species in the tribe Triticeae, and this gene had probably been duplicated before the barley and wheat-rye lineages split, which is variously reported to date back to 8-9 MYA (Middleton et al. 2014) or 11.6 MYA (Martis et al. 2012). Using wheat-barley addition and substitution lines, the chromosomal assignment of the *CENH3* genes in two barley species, *H. vulgare* Linnaeus, 1753 and *H. bulbosum* Linnaeus, 1756, was established as encoding by chromosomes 1H and 6H (Sanei et al. 2011). In hexaploid wheat, *CENH3* genes were assigned to chromosome 1 in all the homeologous genomes, AA, BB and DD (Li et al. 2013, Yuan et al. 2015). However, so far the chromosomal localization of the *CENH3* genes in the other Triticeae species has been beyond the scope of any study known to us.

It has been established by comparative RFLP (restriction fragment length polymorphism) that the rye genome shares extensive synteny with the barley and wheat genomes (Devos et al. 1993). Accumulation of rye genome sequencing data (Martis et al. 2013, Bauer et al. 2017) enabled a genome-wide high-density comparative analysis at a new level and identified six major translocations that had shaped the modern rye genome and made it different from a putative Triticeae ancestral genome (Martis et al. 2013). In this work, we characterize the molecular structure of the main forms of CENH3 protein in rye, *Secale cereale*. Also, we attempt to assign the *CENH3* genes to particular chromosomes, using a set of seven wheat-rye addition lines, each containing a rye chromosome in the wheat genome. A comparison of the localization sites of so functionally important genes in closely related genera is expected to improve our knowledge about conservation

of synteny between the most important cereal species. To this end, available databases with the DNA sequences of rye, *S. cereale*, and diploid species as donors of the hexaploid genome of *T. aestivum* Linnaeus, 1753 have been analyzed and the phylogenetic relationships between the different variants of *CENH3* have been inferred.

Material and methods

Plant material and plant growth

The plant material used were the bread wheat 'Chinese Spring' (CS) (2n=6x=42, AABBDD), the rye cultivar Imperial (2n=2x=14, RR) and wheat-rye ('Chinese Spring'/'Imperial') disomic addition lines involving rye chromosomes 1R–7R (Driscoll and Sears 1971). 'Chinese Spring' is an international standard for wheat research, much as 'Imperial' is for rye. Seeds were from the germplasm collection of the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. All plants were grown in a greenhouse at 22 °C/18 °C (day/night) with a 16-h day length.

RNA isolation and PCR amplification

Total RNA was isolated from leaves of 12-day-old seedlings using the TRI Reagent (MRC, Inc., USA) and treated by RQ-RNase-Free DNase (Promega, Madison, WI) according to the manufacturer's instructions. RNA was reverse-transcribed to cDNA using a RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). The specific primers used to amplify the *CENH3* gene from cDNA were:

1) 5'-ATGGCCCGCACCAAGCAC-3', 5'-GCATCACCAAAGCCTCC-3', to amplify the coding region of αCENH3; and
2) 5'-TGGGTCGCACGAAGCAC-3', 5'-TCACCAAAGCCTTCTCCCC-3', to amplify the coding region of βCENH3.

Cloning and sequencing

RT-PCR products were purified using a Qiagen Purification Kit (Qiagen) and cloned using an InsTAclone PCR Cloning Kit (Thermo Fisher Scientific). Both strands of each of 15–20 clones of each parental variety and addition line were sequenced using an ABI 3130×1 Genetic Analyzer (Applied Biosystems Inc., CA) and an ABI BigDye Kit according to a standard protocol. Similarity searches between the *CENH3* sequences and their orthologs in other species were carried out using the TBLASTN software (Altschul et al. 1990) in the NCBI database (http://blast.ncbi.nlm.nih.gov/Database/).

Sequence alignments and phylogenetic analysis

Amino acids alignments were performed online using Clustal Omega (Sievers et al. 2011) (http://www.ebi.ac.uk/Tools/msa/clustalo) and used for downstream analysis and visualization (Waterhouse et al. 2009) (http://www.jalview.org). The phylogenetic tree based on amino acid sequences was constructed using the Neighbor Joining algorithm in MEGA6 (Tamura et al. 2013). Bootstrap values were calculated from 1000 replicates.

Analysis of databases

The search for rye genomic *CENH3* sequences was performed in among entries in the Sequence Read Archive (European Bioinformatics Institute, accession ID ERP001745) for sorted rye chromosomes 1R-7R (Martis et al. 2013). The sequences of *CENH3* in the putative A-genome donor *Triticum urartu* Thumanjan ex Gandilyan, 1972 (accessions KM507181 and KM507184), the putative B-genome donor *Aegilops speltoides* Tausch, 1837 (accessions KM507182 and KM507185), and the putative D-genome donor *Aegilops tauschii* Cosson, 1849 (accessions KM507183 and KM507186) were downloaded from NCBI. Genomic DNA sequences were aligned with *CENH3*-coding sequences by the Martinez-NW method using Lasergene's MegAlign application.

Results

Assignment of the *CENH3* variants to rye chromosomes

We characterize two main forms of CENH3 proteins, αCENH3 and βCENH3, and their variants, according to differences in size and amino acid substitutions. The *αCENH3-v1* cDNA sequence in the cultivated rye *S. cereale* is 501 bp in length and the associated protein consists of 166 amino acids. In *S. cereale*, *βCENH3-v1* is distinct from *αCENH3-v1* in that the former has several deletions in the NTT and the insertion of three nucleotides, ACC, which encode the amino acid threonine, in the HFD. Thus, *βCENH3-v1* has an overall length of 456 bp and encodes a protein made up by 151 amino acids. Most of the NTT amino acid sequences in αCENH3 and βCENH3 do not align well with each other, the alpha and beta forms share as low as 58% nucleotide identity of the first 60 nucleotides from the 5'-end. In addition to these main forms, their much less common variants (minor, throughout) were found. The *αCENH3-v2* sequences were 492 bp in length each, that is, they were shorter *αCENH3-v1*. Additionally, these two αCENH3 variants have different amino acids at some positions. Some rye accessions carry *CENH3* sequences that are individually 6 bp longer than *βCENH3-v1* and encode two additional amino acid residues of threonine in the NTT domain (Fig. 1B). We classify sequences as these as βCENH3-v2. The two βCENH3 variants have different amino acids at some positions, too.

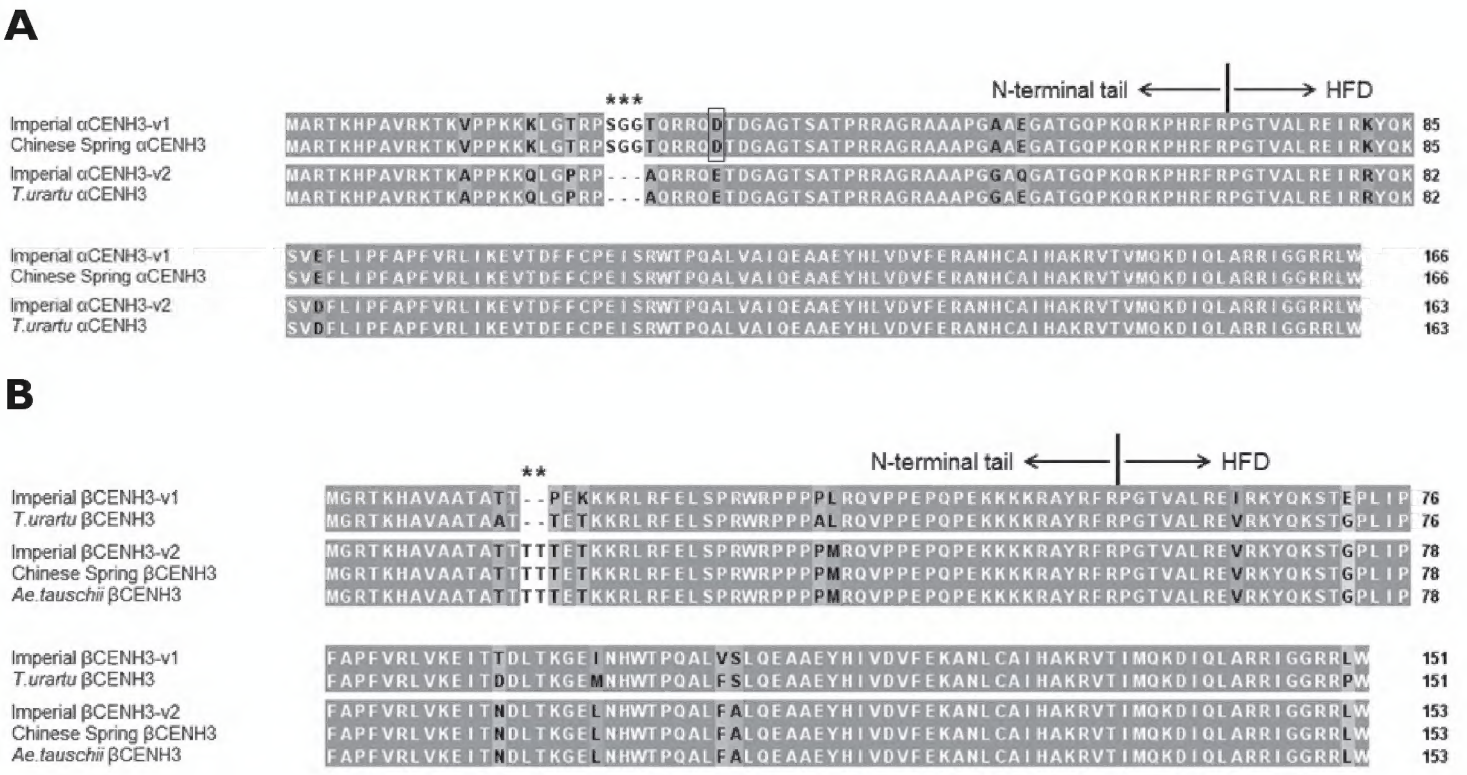

Figure 1. Examples of amino acid alignments of the different variants of the main forms of *CENH3* in rye *S. cereale* (cv. Imperial) and Triticeae species: *T. urartu* (KM507181, KM507184) and *Ae. tauschii* (KM507186). **A** αCENH3 **B** βCENH3. Asterisks are above positions with short insertions/deletions in the N-terminal tail; the position with the highest percentage of amino acid substitutions is framed.

The amino acid differences between CENH3 in rye and wheat were used for the chromosomal assignment of the *CENH3* copies in *S. cereale*. Each of the seven wheat-rye addition lines 'Chinese Spring'/'Imperial' (2n=44 (42+2R)) (Driscoll and Sears 1971) contains the entire hexaploid wheat *Triticum aestivum* genome from 'Chinese Spring' and a pair of rye *S. cereale* chromosomes from 'Imperial'. A comparison of the sequences of the alpha variants of *CENH3* cDNA between 'Imperial' and 'Chinese Spring' revealed a high (99%) level of identity and it is no wonder why the consensus amino acid sequences of 'Imperial' αCENH3-v1 and 'Chinese Spring' αCENH3 were found to be identical (Fig. 1A). The 100% amino acid sequence identity between rye and wheat prevents this variant from being assigned to particular rye chromosomes using addition lines. Several positions in these sequences are polymorphic. The most frequently observed polymorphism is at position 33 of the NTT domain (framed in Fig. 1A), where glutamine acid instead of asparagine acid is present in 20% of the 'Imperial' clones and in 47% of the 'Chinese Spring' clones.

Sixteen percent of the cDNA clones of the alpha variants of 'Imperial' *CENH3* have a 9 bp deletion and represent the minor variant, *αCENH3-v2*, according to our classification (Fig. 1A). None of the 19 'Chinese Spring' clones was found to have this deletion, allowing us to use addition lines for assigning this variant of *CENH3* to rye chromosomes. There was only one addition line (that with rye chromosome 6R), whose clones (39%) had this deletion (Fig. 2). None of the other addition lines had any clone with this deletion.

The shorter, 456-bp-long forms of 'Imperial' *CENH3* DNA produce protein molecules, each containing 156 amino acids and collectively denoted as βCENH3-v1

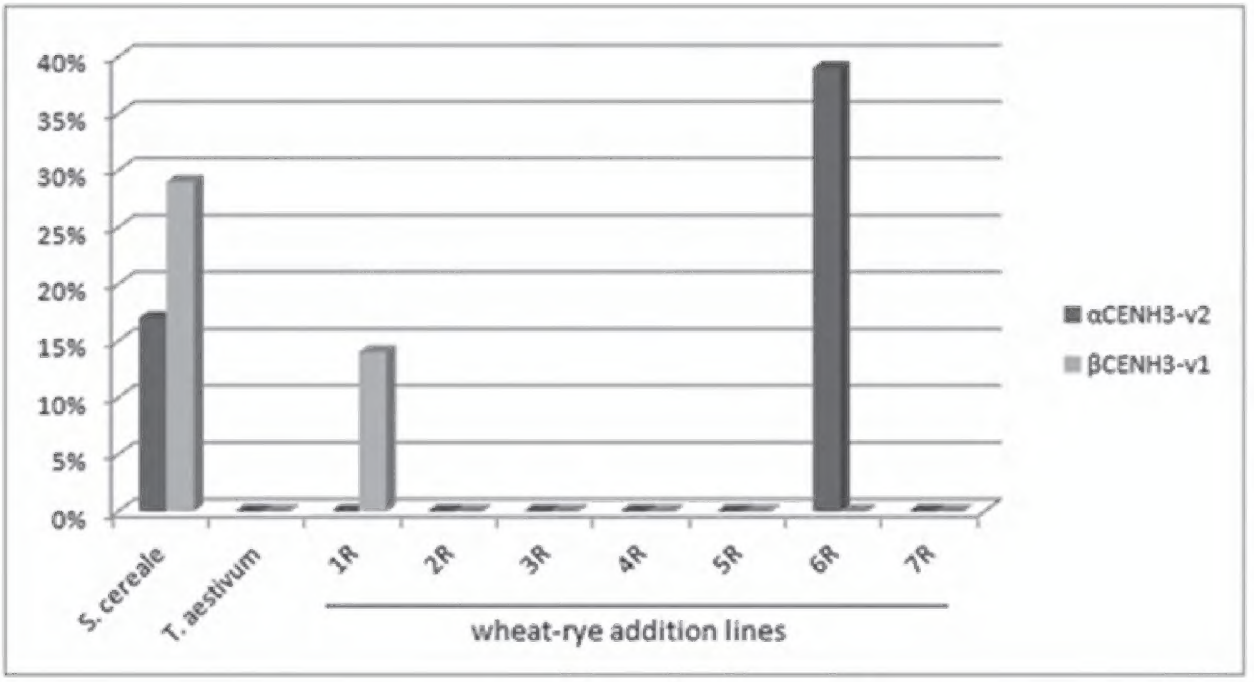

Figure 2. Histogram showing the percentage of α*CENH3*-v2 and β*CENH3*-v1 in the cDNA clones from rye *S. cereale* (cv. Imperial), wheat *T. aestivum* (cv. Chinese Spring) and wheat-rye (1R–7R) addition lines.

(Fig. 1B). They are less common than α*CENH3* and make up 29% of the clones. These sequences did not occur from 'Chinese Spring' DNA amplification and thus their chromosomal assignment using addition lines was possible. There was only one addition line (that with rye chromosome 1R), whose clones (specifically, 14%) were found to have sequences with this structure and of this size (Fig. 2). 'Chinese Spring' cDNA became the source of *CENH3* sequences, each containing a 6-bp insertion in the NTT domain and corresponding, according to our classification, to β*CENH3*-v2 in 'Imperial' DNA. The nucleotide sequence homology of this variant between wheat and rye is 100% and thus addition lines were of as little help in assigning this variant to particular rye as they were with α*CENH3-v1*.

Analysis of databases and phylogenetic relationships between *CENH3* in rye and wheats

To confirm the chromosomal assignment of various *CENH3* variants made using addition lines and to assign α*CENH3*-v1 and β*CENH3*-v2 to particular chromosomes, we analyzed entries in the Sequence Read Archive (European Bioinformatics Institute, accession IDERP001745) for sorted rye chromosomes 1R-7R (Martis et al. 2013). Some sequence reads from chromosome 1R (accession ERX140512) were found to have a high nucleotide identity, 87–97%, of the coding sequence of α*CENH3-v1*. That the contig composed of these reads was really α*CENH3-v1* was indicated by the absence of the 9-bp deletion, typical of α*CENH3-v2*, in one of the reads (ERX140512.2250257), which contained the first exon of α*CENH3* entirely and the coding amino acid sequence identical to αCENH3-v1.

Additionally, two of the reads from chromosome 1R were found to contain the coding region (positions 1 through 328) of β*CENH3-v1* (ERX140512.1955393, ERX140512.290111): they had no 6-bp insertion in the NTT domain that all β*CENH3-v2* normally have and they had large deletions in the NTT that delineate beta forms from alpha forms. Some sequence reads from chromosome 6R (accession ERX140517) were found to have a high nucleotide identity to the HFD in β*CENH3*. Because these reads contained only the most conserved region of the HFD (the last 42 amino acids) and because this region was identical between β*CENH3-v1* and β*CENH3-v2*, we were unable to tell these variants from each other, however, one thing was clear: β*CENH3* is located on chromosome 6R. Thus, the analysis of the Sequence Read Archive for gDNA sequences amplified from sorted rye chromosomes 1R-7R assigned α*CENH3-v1* to chromosome 1R, the beta form of *CENH3* to chromosome 6R and confirmed the addition line-based assignment of β*CENH3-v1* to chromosome 1R. In summary, the main forms of *CENH3* (the alpha and beta forms) are located on rye chromosomes 1R and 6R, the commonest variants, α*CENH3-v1* and β*CENH3-v1*, are on chromosome 1R, and the less common α*CENH3-v2* and probably β*CENH3-v2* are on chromosome 6R. It should be noted that analysis of the most recent version of the rye genome without chromosome sorting (Bauer et al. 2017) did not reveal any *CENH3* variants other than those described herein.

A high level of identity of *CENH3* sequences between wheat and rye is not consistent with a wealth of plant species data that suggest considerable between-species differences in the structure of this protein (Sanei et al. 2011, Masonbrink et al. 2014, Dunemann et al. 2014, Maheshwari et al. 2015). This encouraged us to explore the phylogenetic relationships between the different *CENH3* variants in rye, wheat and diploid species seen as putative donors of the hexaploid wheat genomes. As was found out, *T. urartu* accession KM507181 shares a high (99%) nucleotide identity with 'Imperial' α*CENH3-v2*, and *A. tauschii* accession KM507186, 100% identity with 'Imperial' β*CENH3-v2* (Fig. 1A, B). The neighbor-joining phylogenetic tree for CENH3 amino acid sequences consists of two major clusters, one with alpha forms and another with beta forms (Fig. 3). The topology of the clusters displays some differences in the ways they branch. The bootstrap values in the beta cluster suggest its clearer partitioning into subclusters. In the alpha cluster, CENH3's from *T. urartu* and *A. speltoides* share the same subcluster, while in the beta cluster, each of these species forms a separate branch. CENH3 sequences obtained from the lines, in which added chromosomes (particularly, 1R and 6R) were found to carry rye *CENH3* copies, share the same subcluster with the 'Imperial' sequences. The major variants (variants 1) form prominent subclusters, while the minor variants (variants 2) lie closer to the diploid species *T. urartu, A. tauschii* and *A. speltoides*. Although these species are considered the main contributors to the hexaploid wheat genome, the level of CENH3 identity between them and the rye cultivar Imperial is just as high.

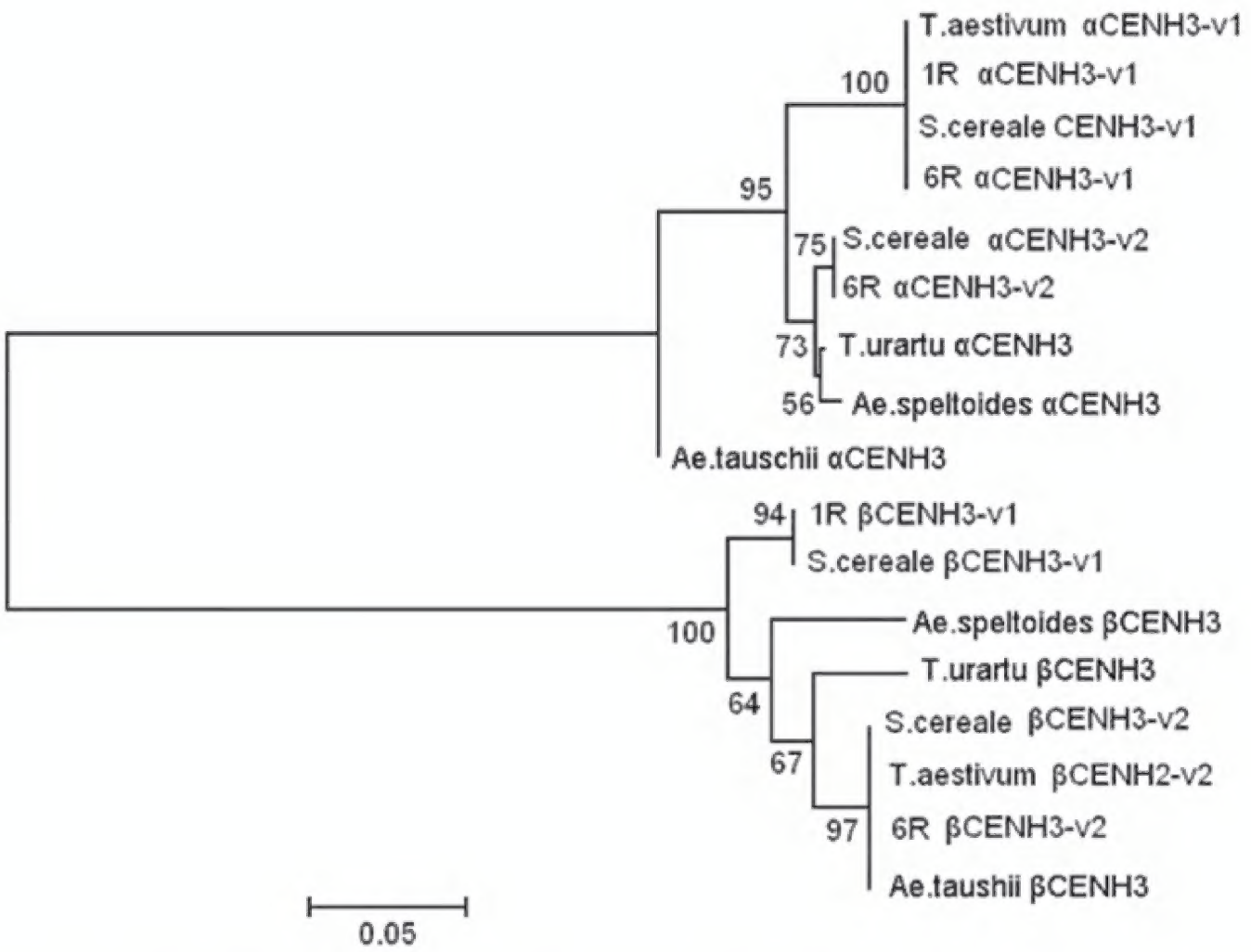

Figure 3. Phylogenetic tree of the CENH3 proteins. Phylogenetic tree inferred using JTT+G models (measures distances) and bootstrapping (1000 replicates). Bootstrap values are indicated on the branches. Rye *S. cereale* (cv. Imperial), wheat *T. aestivum* (cv. Chinese Spring) and wheat-rye (1R–7R) addition lines. NCBI accessions are: αCENH3 and βCENH3 in *T. urartu* (KM507181, KM507184), *Ae. tauschii* (KM507183, KM507186), *Ae. speltoides* (KM507182, KM507185). The scale bar is substitutions per site.

Discussion

Most of the 67 plant species with the CENH3 sequence publicly available – including those that have undergone whole-genome duplication – have this essential gene as a singleton (Maheshwari et al. 2015). Note that sampling bias is a factor here, for 23 species (1/3) were from Brassicaceae. In cereals, three CENH3 variants described in the oats *Avena sativa* Linnaeus, 1753 (Ishii et al. 2015) represent the alpha form and two forms of CENH3 appear for the first time in the closest rye relatives: the Triticeae genera barley, wheat and *Aegilops* (Sanei et al. 2011, Yuan et al. 2015). Comparative analyses of the cereal genomes revealed a whole-genome duplication (WGD), which occurred between 53 and 94 MYA (i.e. before the divergence of the cereal genomes), a recent segmental duplication between chromosomes 11 and 12 and numerous individual gene duplications (Feuillet and Salse 2009). Because the Triticeae species carry a haploid set of seven chromosomes, it is hypothesized that the two forms of CENH3 emerged from a duplication event (Yuan et al. 2015), which in this case is most likely to represent a local duplication event that had taken place before the barley and wheat-rye lineages split.

The now commonly accepted viewpoint authored by Ohno (1970) is that a local duplication event gives rise to two functionally redundant, paralogous genes, one of which will carry on under attenuated selective pressure and will thus be freed from selective constraints. As a result, each copy as this starts to accumulate mutations that would have been deleterious to the gene, if it were a singleton. In reference to *CENH3* in rye, wheat and barley, it can be hypothesized that the copies accumulating deleterious mutations are the beta forms, because their NTT domains contain several deletions of various size that the NTT domains of the alpha forms do not. Interestingly, inactivation of beta CENH3 in barley did not result in a major phenotype and point mutation impairs centromeric *CENH3* loading and induces haploid plants (Karimi-Ashtiyani et al. 2015). The structure of *CENH3* in cereal species that have this gene as a singleton (for example, the rice *O. sativa*) is much closer to the structure of the Triticeae alpha forms (74% identity, while no alignment with the beta form is possible because of extended deletions), which can be regarded as being evolutionarily more ancient than beta forms. According to the theory of evolution by gene duplication (Ohno 1970, Wagner 2002), deleterious mutations should accumulate and lead to a rapid loss of one of the paralogs; however, this is not the case with *CENH3* copies in Triticeae. Among the possible reasons for the observed inconsistency with the theory is the acquisition of new functions by a duplicate gene or by both copies or sub- or neofunctionalization (Conant and Wolfe 2008). For example, in wild tetraploids of wheat, β*CENH3* has a much lower expression level than α*CENH3*, while in cultivated tetraploids β*CENH3* transcripts are enhanced to near α*CENH3* level (Yuan et al. 2015). All the *CENH3* variants that we have described have been found in rye cDNA, that is, they are transcribed; however, to understand the function of each particular variant, special research is required.

According to our results, both main forms, α*CENH3-v1* and β*CENH3-v1*, are located on rye chromosome 1R. Obviously, these copies do not reside there next to each other, but are somewhat spaced. There are two facts that support this statement. First, the two-copy organization of *CENH3* makes these copies very likely to end up with gene conversion events that homogenize their sequences (Wagner 2002). The presence of a large deletion in β*CENH3* and the heterogeneity of the nucleotide and amino acid sequences of the alpha and beta forms suggest quite the opposite. Secondly, analyzing *T. urartu* genome sequencing data (Ling et al. 2013), we found that α*CENH3* and β*CENH3* are 24 kb apart (locus KD187944.1, unplaced genomic scaffold 30245, BioProject PRJNA182347). Considering a high level of identity of *CENH3* sequences between rye and *T. urartu*, it is possible that the genomic region containing *CENH3* has remained highly syntenic between the rye and wheat genomes.

Various molecular mechanisms have been proposed to explain the emergence of these spaced gene copies on the same or different chromosomes (Glover et al. 2015). Some of these mechanisms are 1) retroposition where genes are reverse transcribed and reinserted back into the genome; 2) ectopic recombination during double-strand break (DSB) repair; 3) transposable element mediated gene captured and moved throughout plant genomes; 4) exon shuffling. All these processes involve rearrangements in gene structure. However, the amount of data available so far is insufficient for making an unbiased judgement as to which of these mechanisms should be accepted.

A comparison of the genetic maps between rye and barley shows that rye chromosome 1R and barley chromosome 1H are fully collinear (Martis et al. 2013). Noteworthy, 1R is the only rye chromosome in which no large translocations have been found as the rye karyotype was being shaped from a Triticeae progenitor (Martis et al. 2013). It has previously been established that *CENH3* is located on barley chromosome 1H (Sanei et al. 2011). In wheat, three highly homologous *CENH3* genes have been assigned to chromosomes 1A, 1B and 1D by PCR amplification in nullisomic-tetrasomic lines (Li et al. 2013) and fluorescent in situ hybridization (Yuan et al. 2015). Thus, the presence of *CENH3* genes on chromosome 1 is a shared feature of Triticeae genera. The barley species *H. vulgare* and *H. bulbosum* have additionally been found to have β*CENH3* on chromosome 6H (Sanei et al. 2011). Taking together the existing data and our assignment of *CENH3* copies to rye chromosome 6R, it can be stated that, the *CENH3* genes appeared on chromosome 6 as they did on chromosome 1 in a Triticeae species before the barley and wheat-rye lineages split.

Acknowledgements

This research was financially supported by Russian Fundamental Scientific Research Program on the project 0310-2016-0005, the Russian Foundation for Basic Research (grant 17-04-00748a). The bioinformatics analysis was undertaken with the support of the ICG SB RAS budget project VI.61.1.2.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Bauer E, Schmutzer T, Barilar I, Mascher M, Gundlach H, Martis MM, Twardziok SO, Hackauf B, Gordillo A, Wilde P, Schmidt M, Korzun V, Mayer KFX, Schmid K, Schon C-C, Scholz U (2017) Towards a whole-genome sequence for rye (*Secale cereale* L.). The Plant Journal 89: 853–869. https://doi.org/10.1111/tpj.13436

Conant GC, Wolfe KH (2008) Turning a hobby into a job: How duplicated genes find new functions. Nature Review Genetics 9: 938–950. https://doi.org/10.1038/nrg2482

Cooper JL, Henikoff S (2004) Adaptive evolution of the histone fold domain in centromeric histone. Molecular Biology and Evolution 21: 1712–1718. https://doi.org/10.1093/molbev/msh179

Devos KM, Atkinson MD, Chinoy CN, Francis HA, Harcourt RL, Koebner RMD, Liu CJ, Masojc P, Xie DX, Gale MD (1993) Chromosomal rearrangements in the rye genome relative to that of wheat. Theoretical and Applied Genetics 85: 673–680. https://doi.org/10.1007/BF00225004

Driscoll CS, Sears ER (1971) Individual addition of the chromosomes of Imperial rye to wheat. Agronomy Abstract 1971: 6

Dunemann F, Schrader O, Budahn H, Houben A (2014) Characterization of centromeric histone H3 (CENH3) variants in cultivated and wild carrots (Daucus sp.). PLoS ONE 9(6): e98504. https://doi.org/10.1371/journal.pone.0098504

Feuillet C, Salse J (2009) Comparative Genomics in the Triticeae. In: Feuillet C, Muehlbauer GJ (Eds) Genetics and Genomics of the Triticeae, Plant Genetics and Genomics: Crops and Models 7. Springer Science+Business Media, 451–477. https://doi.org/10.1007/978-0-387-77489-3_17

Finseth FR, Dong Y, Saunders A, Fishman L (2015) Duplication and adaptive evolution of a key centromeric protein in *Mimulus*, a genus with female meiotic drive. Molecular Biology and Evolution 32: 2694–2706. https://doi.org/10.1093/molbev/msv145

Glover NM, Daron J, Pingault L, van de Poele K, Paux E, Feuillet C, Choulet F (2015) Small-scale gene duplications played a major role in the recent evolution of wheat chromosome 3B. Genome Biology 16: 188. https://doi.org/10.1186/s13059-015-0754-6

Hirsch CD, Wu Y, Yan H, Jiang J (2009) Lineage-specific adaptive evolution of the centromeric protein CENH3 in diploid and allotetraploid *Oryza* species. Molecular Biology and Evolution 26: 2877–2885. https://doi.org/10.1093/molbev/msp208

Ishii T, Karimi-Ashtiyani R, Banaei-Moghaddam AM, Schubert V, Fuchs J, Houben A (2015) The differential loading of two barley CENH3 variants into distinct centromeric substructures is cell type and development-specific. Chromosome Research 23: 277–284. https://doi.org/10.1007/s10577-015-9466-8

Karimi-Ashtiyani R, Ishii T, Niessen M, Stein N, Heckmann S, Gurushidze M, Banaei-Moghaddam AM, Fuchs J, Schubert V, Koch K, Weiss O, Demidov D, Schmidt K, Kumlehn J, Houben A (2015) Point mutation impairs centromeric CENH3 loading and induces haploid plants. Proceedings of the National Academy of Sciences of the USA 112: 11211–11216. www.pnas.org/cgi/doi/10.1073/pnas.1504333112

Li b, Choulet F, Heng Y, Hao W, Paux E, Liu Z, Yue W, Jin W, Feuillet C, Zhang X (2013) Wheat centromeric retrotransposons: the new ones take a major role in centromeric structure. The Plant Journal 73: 952–965. https://doi.org/10.1111/tpj.12086

Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y et al. (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496: 87–90. https://doi.org/10.1038/nature11997

Maheshwari S, Tan EH, West A, Franklin FCH, Comai L, Chan SWL (2015) Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. PLoS Genetics 11 (1): e1004970. https://doi.org/10.1371/journal.pgen.1004970

Malik HS, Henikoff S (2001) Adaptive evolution of Cid, a centromere-specific histone in Drosophila. Genetics 157: 1293–1298

Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Simkova H, Novak P, Neumann P, Kubalakova M, Bauer E, Haseneyer G, Fuchs J, Dolezel J, Stein N, Mayer KFX, Houben A (2012) Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. Proceedings of the National Academy of Sciences of the USA 109: 13343–13346. www.pnas.org/cgi/doi/10.1073/pnas.1204237109

Martis MM, Zhou R, Haseneyer G, Schmutzer T, Vrana J, Kubalakova M, Konig S, Kugler KG, Scholz U, Hackauf B, Korzun V, Schon C-C, Dolezel J, Bauer E, Mayer KFX, Stein N (2013) Reticulate Evolution of the Rye Genome. The Plant Cell 25: 3685–3698. www.plantcell.org/cgi/doi/10.1105/tpc.113.114553

Masonbrink RE, Gallagher JP, Jareczek JJ, Renny-Byfield S, Grover CE, Gong L, Wendel JF (2014) CenH3 evolution in diploids and polyploids of three angiosperm genera. BMC Plant Biology 14: 383. https://doi.org/10.1186/s12870-014-0383-3

Middleton CP, Senerchia N, Stein N, Akhunov ED, Keller B, Wicker T, Kilian B (2014) Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the Triticeae tribe. PLoS ONE 9(3): e85761. https://doi.org/10.1371/journal.pone.0085761

Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J (2004) Sequencing of a rice centromere uncovers active genes. Nature Genetics 36: 138–145. https://doi.org/10.1038/ng1289

Neumann P, Pavlíková Z, Koblížková A, Fuková I, Jedličková V, Novák P, Macas J (2015) Centromeres off the hook: massive changes in centromere size and structure following duplication of *CenH3* gene in *Fabeae* species. Molecular Biology and Evolution 32: 1862–1879. https://doi.org/10.1093/molbev/msp070

Ohno S (1970) Evolution by gene duplication. Springer Science+Business Media, New York. https://doi.org./10.1007/978-3-642-86659-3

Sanei M, Pickering R, Kumke K, Nasuda S, Houben A (2011) Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proceedings of the National Academy of Sciences of the USA 108: 498–505. www.pnas.org/cgi/doi/10.1073/pnas.1103190108

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7: 539. https://doi.org/10.1038/msb.2011.75

Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S (2002) Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. The Plant Cell 14: 1053–1066. https://doi.org/10.1105/tpc.010425

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30: 2725–2729. https://doi.org/10.1093/molbev/mst197

Vermaak D, Hayden HS, Henikoff S (2002) Centromere targeting element within the histone fold domain of Cid. Molecular and Cellular Biology 22: 7553–7561. https://doi.org/10.1128/MCB.22.21.7553–7561.2002

Wagner A (2002) Selection and gene duplication: a view from the genome. Genome Biology 3(5): reviews 1012.1–1012.3. https://genomebiology.com/2002/3/5/reviews/1012

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. Bioinformatics 25: 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Yuan J, Guo X, Hu J, Lv Z, Han F (2015) Characterization of two CENH3 genes and their roles in wheat evolution. New Phytologist 206: 839–851. https://doi.org/10.1111/nph.13235

Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK (2002) Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. The Plant Cell 14: 2825–2836. https://doi.org/10.1105/tpc.006106